

**AMENDMENTS TO THE CLAIMS**

This listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims:**

1-17. (Canceled).

18. (Currently Amended)     A method of operating a video surveillance system, the video surveillance system including at least one sensing unit capable of being operated in first and second modes, the method comprising:

operating a sensing unit in the first mode to automatically scan a scene for targets;

processing image data of the scene from the sensing unit in the first mode to automatically detect the presence of an interesting target in the scene;

upon detecting the interesting target, operating the sensing unit in the second mode to automatically track the interesting target and to obtain higher quality imagery of a subset of the scene including the interesting target than that obtained by operating the sensing unit in the first mode by performing at least one of: panning, tilting, and zooming, using super-resolution algorithms, digitally zooming into the image, employing different lenses or filters, or using a robotic actuator to obtain the higher quality imagery of the interesting target than that obtained by operating the sensing unit in the first mode; and

processing image data from the sensing unit in the second mode to automatically track the target by sending at least one of pan, tilt, and zoom commands to the sensing unit.

19-20. (Canceled).

21. (Previously presented) The method of Claim 18, wherein processing image data from the sensing unit in a first mode comprises :

processing the image data with a vision module to detect and locate at least one target by at least one of:

- 1) geo-locating the at least one object in 3D space; or
- 2) classifying pixels in the image information as background pixels or foreground pixels, using the foreground pixels to determine at least one blob, tracking at least one possible target based on at least one blob, determining when at least one blob merges or splits into one or more possible targets, and performing at least one of:
  - a) filtering and predicting the image location of at least one of the possible targets, or
  - b) calculating a 3D position of at least one of the possible targets if calibration information is available;

determining if at least one predetermined condition has been violated by at least one target;  
and

classifying at least one possible target by target type into categories.

22-26. (Canceled).

27. (Previously Presented) The method of Claim 21, further comprising providing summary statistics including at least one of the following: an estimate of the target's position in image space; the target's position in a world coordinate frame; the target's size in image space; the target's size in the world coordinate frame; the target's velocity in image space; the target's velocity in the world coordinate frame; the target's trajectory; the target's color; the color of at least one subcomponent of the target; a bounding box for the target in an image; or at least one estimate of location or size of at least one body part, in the case where the target is human to be used in determining if at least one predetermined condition has been violated by at least one object.

28. (Canceled).

29. (Previously Presented) The method of Claim 18, wherein processing image data from the sensing unit in the second mode, if the second mode involves camera motion, comprises:

segmenting the target from the background by detecting motion of the sensing unit and factoring it out;

calculating color histogram information in a region of at least some pixels of image data from the sensing unit;

calculating an edge image; and

using template matching, to integrate the results of the steps of motion segmentation, edge detection and color histogram formation, to obtain an estimate of the target's position.

30. (Previously Presented) The method of Claim 29, wherein processing image data from the sensing unit in a second mode further comprises:

generating summary statistics including at least one of an estimate of the target's position in image space; the target's position in a world coordinate frame; the target's size in image space; the target's size in the world coordinate frame; the target's velocity in image space; the target's velocity in the world coordinate frame; the target's trajectory; the target's color; the color of at least one subcomponent of the target; a bounding box for the target in an image; or at least one estimate of location or size of at least one body part, in the case where the target is human, based on various features of the target.

31. (Canceled).

32. (Previously Presented) The method of Claim 29, wherein processing image data from the sensing unit in a second mode further comprises:

monitoring a length of time in which the system has been in the second mode;

monitoring the various summary statistics to detect if any target characteristic has met a predetermined rule for switching to the first mode;

monitoring for any external stimuli indicating that the system should switch to the first mode; and

if any predetermined conditions are met in any of the monitoring processes, switching back to the first mode.

33. (Previously Presented) The method of Claim 18, further comprising:

determining a best shot of the target while operating the visual module in either of the first and second modes; and

delivering the best shot back to the user through the response module.

34. (Previously Presented) The method of Claim 18, further comprising: if the target exits the field of view of the sensing unit,

receiving, by a second sensing unit, information about a last known position of the target from the sensing unit;

operating the second sensing unit in the first mode to scan for the target, focusing on a region near the last known position;

processing image data from the second sensing unit in the first mode to detect the presence of a target;

upon detecting a target, operating the second sensing unit in the second mode to track the target; and

processing image data from the second sensing unit in the second mode to track the target by sending at least one of pan, tilt, and zoom commands to the second sensing unit,

wherein, if the target is not detected within a predetermined time period in the step of operating the second sensing unit in the first mode to scan for the target, focusing on a region near the last known position, the second sensing unit switches to scanning for targets in its entire field of coverage.

35-37. (Canceled).

38. (Original) The method of Claim 18, wherein, if operating the sensing unit in the second mode involves camera motion, processing image data from the sensing unit in a second mode comprises:

- (a) performing corner detection to find interesting points;
- (b) searching for matches for interesting points in successive images;
- (c) assigning confidence values to the matches of the interesting points;
- (d) employing a robust averaging method to determine an estimate from a subset of the matches determined to have high confidence values;
- (e) warping a second successive image towards a first successive image; and
- (f) subtracting the warped image from a current image to determine which pixels have moved.

39. (Previously Presented) The method of Claim 38, wherein processing image data from the sensing unit in a second mode further comprises:

- performing (a)-(f) on reduced resolution images; and

performing (a)-(f) on full resolution images, utilizing the results of performing (a)-(f) on reduced resolution images.

40. (Previously Presented) The method of Claim 38, wherein processing image data from the sensing unit in a second mode further comprises:

using three-frame differencing.

41-62. (Canceled).

63. (Previously Presented) The method of claim 18, further comprising at least one of the following: sending an e-mail alert; sounding an audio alarm; providing a visual alarm; logging an alert to a persistent medium; making a telephone call; transmitting a wireless message to a personal digital assistant; enabling a mechanical actuator; or providing position information to another sensing unit.

64. (Currently Amended) A video surveillance system, comprising:

at least one sensing unit capable of being operated in first and second modes, the sensing unit operating in the first mode to automatically scan a scene for a target ;

a response module for, upon detecting the target, operating the sensing unit in the second mode to automatically track the target and obtain higher quality imagery of a subset of the scene including the interesting target than that obtained by operating the sensing unit in the first mode by

performing at least one of: panning, tilting, and zooming, using super-resolution algorithms, digitally zooming into the image, employing different lenses or filters, or using a robotic actuator to obtain the higher quality imagery of the interesting target than that obtained by operating the sensing unit in the first mode; and

a computer system coupled to the sensing unit, the computer system receiving and processing image data from the sensing unit, automatically detecting and tracking targets, and determining whether the sensing unit operates in the first mode or in the second mode based on the detection and tracking of targets.

65. (Previously Presented) The method of Claim 64, wherein the computer system further comprises:

a vision module processing the image data to detect and locate at least one target by at least one of:

- 1) geo-locating the at least one object in 3D space; or
- 2) classifying pixels in the image information as background pixels or foreground pixels, using the foreground pixels to determine at least one blob, tracking at least one possible target based on at least one blob, determining when at least one blob merges or splits into one or more possible targets, and performing at least one of:

- a) filtering and predicting the image location of at least one of the possible targets, or



b) calculating a 3D position of at least one of the possible targets if calibration information is available;  
determining if at least one predetermined condition has been violated by at least one target;  
and  
classifying at least one possible target by target type into categories.

66. (Previously Presented) The method of Claim 65, wherein the vision module further provides summary statistics including at least one of the following: an estimate of the target's position in image space; the target's position in a world coordinate frame; the target's size in image space; the target's size in the world coordinate frame; the target's velocity in image space; the target's velocity in the world coordinate frame; the target's trajectory; the target's color; the color of at least one subcomponent of the target; a bounding box for the target in an image; or at least one estimate of location or size of at least one body part, in the case where the target is human to be used in determining if at least one predetermined condition has been violated by at least one object.

67. (Previously Presented) The method of Claim 64, wherein the vision module further comprises:

a segmentation module to separate background pixels from foreground pixels;  
a blobizer to receive the foreground pixels from the segmentation module and to group connected foreground pixels into coherent blobs;

a target tracker to process the coherent blobs, determine when they are targets, and to obtain position information for each target;

a classifier to determine a target type for each target; and

a primitive generation module to generate summary statistics to be sent to the inference module.

68. (Previously Presented) The method of Claim 67, wherein the vision module generates summary statistics including at least one of an estimate of the target's position in image space; the target's position in a world coordinate frame; the target's size in image space; the target's size in the world coordinate frame; the target's velocity in image space; the target's velocity in the world coordinate frame; the target's trajectory; the target's color; the color of at least one subcomponent of the target; a bounding box for the target in an image; or at least one estimate of location or size of at least one body part, in the case where the target is human, based on various features of the target.

69. (Previously Presented) The method of Claim 67, further comprising an inference module which, in the second mode, determines whether to switch back to the first mode by monitoring at least one of how long the system has been in the second mode, the position of the target in the image, the position of the target in a world coordinate frame, and one or more external stimuli.

70. (Previously Presented) The method of Claim 64, wherein the vision module determines a best shot of the target while operating the vision module in either of the first and second modes; and

a response module coupled to the vision module, the vision module delivering the best shot back to the user through the response module.

71. (Previously Presented) The method of Claim 64, further comprising:

a second sensing unit coupled to the sensing unit to receive information about a last known position of the target from the sensing unit; if the target exits the field of view of the sensing unit, the second sensing unit operating in the first mode to scan for the target, focusing on a region near the last known position and operating in the second mode to track the target upon detecting the target; and;

the second sending unit coupled to the computer system, the computer system processing image data from the second sensing unit in the first mode to detect the presence of a target, processing image data from the second sensing unit in the second mode to track the target by sending at least one of pan, tilt, and zoom commands to the second sensing unit,

wherein, if the target is not detected within a predetermined time period, the second sensing unit operates in the first mode to scan for the target, focusing on a region near the last known position, and switches to scanning for targets in its entire field of coverage.

72. (Currently Amended) A computer readable medium for use in a video surveillance system including at least one sensing unit capable of being operated in first and second modes, the computer-readable information storage medium storing computer-readable program code for causing the computer to perform the steps of:

processing image data from the sensing unit operating in the first mode to detect the presence of an interesting target in a scene;

upon detecting the interesting target, operating the sensing unit in the second mode to automatically track the interesting target and to obtain higher quality imagery of a subset of the scene including the interesting target than that obtained by operating the sensing unit in the first mode by performing at least one of: panning, tilting, and zooming, using super-resolution algorithms, digitally zooming into the image, employing different lenses or filters, or using a robotic actuator to obtain the higher quality imagery of the interesting target than that obtained by operating the sensing unit in the first mode; and

processing image data from the sensing unit in the second mode to automatically track the target by sending at least one of pan, tilt, and zoom commands to the sensing unit.